DOCUMENT RESUME

ED 243 793 SO 015 645

AUTHOR Geis, George L.; And Others
TITLE Instructional Development: Developmental Testing.
PUB DATE 84
NOTE 66p.; Paper presented at the Annual Meeting of the American Educational Research Association (68th, New Orleans, LA, April 23-27, 1984). For document on revision procedures, see SO 015 646.
PUB TYPE Information Analyses (070) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *Curriculum Development; Curriculum Evaluation; Curriculum Research; Educational Assessment; Field Tests; *Formative Evaluation; Higher Education; *Instructional Design; Instructional Development; Material Development
IDENTIFIERS *Revision Processes

ABSTRACT

A review of the research literature on systematic instructional design indicates that few empirically supported guidelines exist for testing materials under development and for improving materials on the basis of feedback from potential learners. In an analysis of 17 textbooks on instructional design, eight specific aspects to be considered in developmental testing of instruction materials were identified: criterion performance, number of students, population characteristics, role of students, role of developer, method of data collection, basis of revision, and testing environment. Findings indicate: (1) some, but not universal, agreement on particular aspects of developmental testing, (2) major differences concerning the nature of developmental testing activities, (3) a plethora of general advice but few specific guidelines, (4) little linkage between guidelines for developmental testing and guidelines for revision, (5) no apparent recognition of and no reference to other areas of education and psychology which might contribute to a technology of developmental testing, and (6) little research support for the proposed guidelines. Appendices include a bibliography, flow charts, a summary table noting each author's suggestions regarding each of the eight categories of developmental testing, and a survey questionnaire on the process of developmental testing and revision of prototype instructional materials. (LH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED243793

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.
X Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
George L. Geis
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

INSTRUCTIONAL DEVELOPMENT: DEVELOPMENTAL TESTING[1]

George L. Geis

Charles Burt

Cynthia Weston

McGill University

Paper presented at the Annual Meeting of the

American Educational Research Association

(68th, New Orleans, LA, April 23-27, 1984)

SO 015 645

Instructional Development: Developmental Testing[1]

George L. Geis

Charles Burt

Cynthia Weston

McGill University

An outstanding attribute of systematic instructional design (as described, e.g., by Davis,et al,1974;Deterline and Lenn,1972,) is its data-based nature. For example, what is taught should be determined by needs assessment and performance analysis rather than a priori by a given content. And, figuring prominently in almost all descriptions of the instructional development process are the steps of "tryout and revision." Andrews and Goodson (1980) found, in 34 of the 40 models of systematic instructional design which they examined,the iterative cycle of testing and revision was present. Clearly there is a commitment among instructional designers to testing out materials under development and improving them on the basis of feedback from potential learners. This type of formative evaluation has been labelled Developmental Testing.(See Note 2.)

Developmental testing is so central to the design process that one might expect the literature would provide well worked out guidelines or algorithms for it and for revision. This paper and its companion (Weston, et al,1984) describe the results of our search for such guidelines.

Applying front end analysis, another key concept in systematic

instructional design, one should begin by asking: What are the procedures for revision? Then: What kind of data are required for carrying out the revision procedures? And, finally: What activities during developmental testing will produce these useful data? Unfortunately, as our paper on revision shows, there is no single algorithm (or set of algorithms, or heuristics) for revision. Consequently, it is not possible to work backwards toward, or to construct, the procedures that would produce the data required for effective revision. Nevertheless, we can report on the advice, suggestions, and guidelines for developmental testing which appear in the literature.

During the past year we have conducted several literature searches. The ERIC collection was explored. We examined recent volumes of a number of journals. We reviewed about 30 textbooks on instructional design from which we selected 17 which devote substantial space to developmental testing; we analyzed the relevant sections in detail. (The references for the selected sample may be found in Appendix A.) We will report here primarily the information obtained from the textbooks with occassional references to key journal articles or a dissertation.

Although there is a limited amount of material on developmental testing and revision, a few authors have contributed extensively to the area. For example, Robert Horn (1966) offers a self-instructional training package for novices. George Gropper (1975) provides a detailed analysis of and recommendations for diagnosis and revision . Allan Abedor (1971) offers a review of existing models and presents a validated model of his own for development testing and revision. Walter Dick, as well as providing developmental testing guidelines in the textbook he co-authored (Dick and

Carey,1978) has regularly monitored the current status of research on developmental testing and reported his findings (e.g.,Dick,1980.) Tom Lawson (1974) offers essentially a collection of "job aids" for the developmental tester,and includes recognition of the work of Sanders and Cunningham (1973.)

Each suggestion or guideline that appeared in the textbooks was added to a pool of statements. The pool was then examined for general areas or topics. From this informal content analysis we constructed a set of eight categories (listed in Appendix B.) This enabled us to construct the summary table (Table I) which briefly notes each author's suggestions regarding the eight specific aspects of developmental testing and revision.

First,some observations concerning the categories and the suggestions in them.

Criterion Performance. A number of authors have addressed the critical question: when is instruction good enough? The answer, of course, implies the point at which the designer may terminate the iterative cycle of developmental testing and revision.

Generally,criterion performance is defined in terms of post-instruction test data. For example, a score of 70-80% on the posttest is often suggested as evidence of satisfactory performance of the preceeding instruction. Some authors suggest that no one item on the posttest should be failed by more than 20% of the population. (This criterion is a bit more lenient than the one proposed in training area of 90/90 ---90% of the students score 90% or more on the posttest.)

This emphasis on post-test scores assumes that the test itself is technically sound and that it comprehensively samples the skill and knowledge domains represented in the teaching materials, not an insignificant consideration. Furthermore, as discussed elsewhere (Weston, et al 1984,) authors regularly suggest that posttest items be specifically keyed to particular segments of the instruction. This permits appropriate diagnosis activities to be carried out.

Some authors suggest a criterion based on errors within the instructional material itself. This is typically referred to as "frame error analysis" recalling the historical importance of the programmed instruction movement of the 1960's in the impetus for developmental testing. Questions imbedded in the text or probe questions presented by the tester might equally well produce such error data.

Finally, some authors add a criterion which refers to learner reactions---attitudinal data. For example, the results of a reactionnaire should show that no more than 20% of the students indicate dissatisfaction with the teaching materials.

As was noted earlier, the activity of developmental testing achieved prominance during the period when programmed instruction was flourishing. Hence the suggestions for testing procedures (and for revision) often assume that the teaching materials are in a frame-by-frame format. Microscopic examination of student performance during learning (e.g., error analysis of each frame) supposes an instructional or testing format which provides regular opportunities for overt student response.

Number of students. The size of the suggested test population depends very much upon the concept of the developmental testing session(s) held by the author. Thus: if testing is seen as an intensive,clinical, probing activity, the author is likely to suggest one-on-one sessions (or at most using two or three students at the same time.) On the other hand, small groups (e.g., six subjects) are suggested by authors who opt for emphasis on data gathering during the instructional session and greater emphasis on posttest data.

The answer to the question of how many test subjects should be used is also related to the profile of the test sample. The author who suggests one-on-one sesssions is likely to also suggest repetitions of such sessions using different learners until:(a) learning criteria are reached and (b) an appropriate sample of the target population has been covered.

Population. There is general agreement that the test group should be representative of the target population. (Although Johnson and Johnson ,1975) propose that anyone can serve as a test subject, presumably they expect the developer to exercise common sense in the selection.) Less clear are the answers to the questions: How similar should the sample population be to the target population? And, along which dimensions should the two populations coincide? Some authors suggest merely choosing "high, low, and average learners;" some more specifically suggest using pretest data or other pre-instructional information such as SAT scores to aid in selection. Lawson (1974) opts for a poorer learner on the assumption that if poorer learners can work their way through the materials successfully, all others will be able to do so. A few emphasize personal learner characteristics such as "the test subject should not be timid or docile." Presumably more articulate and

aggressive students are likely to volunteer comments and criticisms, an especially important consideration for those engaged in one-on-one testing.

Role of the student. How an author generally conceives of developmental testsing is best illustrated by his/her descriptions of student and developer roles.

The variety of proposed student roles can be visualized as in Figure 1. At the one extreme the test subject is seen as a traditional student: he/she acquires the appropriate knowledge and skills to the degree indicated by the results of the posttest. The primary data upon which the developer bases decisions (e.g., continue testing,make revisions),then, are posttest results. At the other extreme is the test subject who is encouraged to think aloud, make comments and suggest revisions while working through the learning materials. In between are roles which range from occasional comments by the learner and limited questionning by the developer through much more active commenting, and critiquing by the test subject. In a number of cases this more active role of the student is prescribed for a post-instructional debriefing session ,involving an individual or a group.

Comments by learners are in some cases to be written by the learner (e.g., on a comment sheet, on the learning materials), or they may be recorded by the observer-developer (e.g., on a master copy of the instructional material or using a tape recorder.)

At the outset of the testing session students may be instructed in various ways, e.g., "Go through the instruction as best you can." "Ask

questions only when you are really stuck." "Comment on points of difficulty as you work through the materials."

In summary, suggested student roles range from that of test-subject-as-learner to test-subject-as-critic.

One important point already noted is the relation between degree of student activity during testing and the emphasis on the posttest. Obviously, the more that the student interacts with the tester (and as a consequence elicits explanations, on-the-spot revisions,etc.) the less are the posttest results a reflection of the original instructional materials. Consequently it is essential that developers using a highly intervening strategy must maintain precise descriptions of changes made during testing (see: Role of the Developer,below.)

Role of the developer. Like that of the test subject, the role of the developer is described quite differently by various authors. Again a continuum may be visualized (Figure 2.)

Suggestions for developer role during testing range from that of a testing session supervisor (e.g., gives instructions for the session, records learning time) through the passive observer (e.g.,records student's reactions) to active intervener (either as a tutor or as someone who probes for comments and asks for suggestions. In some cases he/she offers posssible revisions on the spot.) Degree of intervention varies quite a bit among authors; some suggest a rather reactive role---responding succinctly to student questions;others propose that the developer intervene regularly,actively probing and encouraging comments.

There is general agreement that in the post-instruction debriefing sessions the developer should set the agenda, probe for difficulties learners encountered, and summarize comments. (Specific guidelines for conducting debriefing sessions and for probing during instruction are offered by Gropper,1975,pps 97-101.)

We should note that we use the term "developer" as if one person were carrying out all the activities. There are some authors who recommend a differention of roles, for example that the person who gathers the data during developmental testing should not be the author of the materials.

Data collection. Every author suggests that a posttest be employed, although, as already noted, authors assign quite different emphases to the data it provides. And most of them suggest that in addition some kind of attitudinal questionnaire be employed.

Those authors who assign to the test subject an active role during instruction propose various means of collecting data during the test session: tape recording, written student comments, written observations by the observer, proposed changes noted on the master copy of the instructional material, "frame" error data, notes on assistance required during learning (i.e.,tutoring), notes on the non-verbal behavior of the learner,etc.

Testing environment. Finally, we uncovered a few (surprisingly few) comments of the actual physical environment during testing and debriefing. Only Robert Horn (1966) seems to have seriously considered the physical environment and, in general, his suggestions seem like obvious reminders. Others (e.g.,Abedor 1971, Paulson ,1969) have incorporated his suggestions

with their own. More attention might well be paid to the testing conditions since they may represent a not trivial consideration in the training of novice instructional developers.

Summary. Some overall observations can be made.

1.There is some but not universal agreement on particular aspects of developmental testing. For example: the use of criterion or post-instruction tests, sample representative learners, the use of materials in rough draft rather than finished form.

2. There are major differences concerning both the overall nature of developmental testing activities and particular features of these activities (e.g., role of the student, role of the developer.)

3. As the degree of precision of the advice offered is increased the number of authors offering the advice decreases (see Figure 3.) In other words, there is a plethora of general advice but very few algorithms.

4.There is little linkage between guidelines for developmental testing and guidelines for revision. Especially notable is the lack of discussion of how the particular kinds of data obtained during developmental testing will be (or can be) used in revising the instruction. As Dick (1980) says:

> "The greatest limitation in formative evaluation today (is) the dilemma of what to do after a problem has been detected in instruction. Nearly all instructional writers have indicated that after the data have been collected and summarized in a formative evaluation, the designer should 'revise appropriately'. However, in most instances, designers have already used their best knowledge of how to design the instruction, and therefore it is not always apparent what 'revising appropriately' would be." (p.5)

5. There is apparently no recognition of and no reference to other areas in Education and Psychology which might contribute to a technology of developmental testing. For example, Social and Clinical Psychology and Counselling might offer extensive information about interviewing techniques which would allow instructional developers to elaborate on the sketchy suggestions in the literature on developmental testing. Cognitive psychologists, as they microscopically examine the procedures learners use as they acquire and recall new knowledge and skills, have accumulated research data and lore on such matters as collecting and analyzing learner protocols. Furthermore,this area of Psychology offers insights into techniques of error analysis and even revision procedures which might well be relevant to those engaged in developmental testing and revision. Knowledge of major advances in testing, especially criterion-referenced testing, should clearly be of assistance to the developer.

6.There is almost no research support or even clear rationales for most of the guidelines proposed.

Research. Let us pause on this last point for a moment.

Periodically reviews of research in developmental testing appear in the literature (e.g., Baker and Alkin 1973, Dick 1980.)

Recently,an issue of one journal has been devoted to the area and includes resea ch reports (Performance and Instruction, 1983.)

There is general recognition that research in this area is difficult. That is perhaps to be expected in an any area of applied,technological

research. It must pass through the Scylla and Charybdis of, on the one side, too little control over a multitude of variables in realistic environments, and on the other side, a well-controlled but highly artificial environment. In the first case the researcher is unlikely to be able to demonstrate the effects of particular variables though the results do accurately reflect the impact of a real environment. In the second case the control that is exercised increases the chances of identifying critical variables but generalization to the more complex real environment is sorely limited.

The inherent difficulties in carrying out applied research that is technically sound and directly applicable to the real world probably explains in part the paucity of research data in developmental testing.

Nevertheless there is a worthwhile and suggestive set of studies that have been carried out over the past few years. For example,the number and use of test subjects has been examined by Robeck (1965) and Abedor (1971.) Some attempts have been made to test the effectiveness of heuristics for developmental testing ,e.g., Robinson (1972),Abedor (1971.)

More recently, Wager (1983) compared the results of one-to-one and small group testing, taking into account student aptitudes. She concludes that by paying careful attention to the mix of student aptitudes the developer may avoid having to run both one-to-one and group tests.

Many areas within developmental testing deserve research study. Dick (1980) proposes three broad concerns that might be investigated:

—organization and administration,i.e., "who does what to whom in the

formative evaluation process ? " (p.4.)

---"general applicability of the formative evaluation process to a wide range of instructional design settings."(p. 5)

---"a theory-based approach to formative evaluation."(p. 5)

On this last point, it should be pointed out that,in a real sense, developmental testing and revison procedures are continually testing out the "theory" of design from which the developer draws his/her procedures.

Can precise guidelines be constructed for developmental testing? Of course, one way to resolve this question is to declare it a non-problem. Thus, one could assert either that there can never be algorithms for developmental testing or even that such guidelines are not necessary.

On the first point, Susan Markle, a long-time, insightful contributor to instructional technology, has stated:

> "Procedures for eliciting (developmental testing) data vary. Some testers prefer to talk to the students throughout the process... Others prefer to query the student who hesitates or errs, leaving him to his own devices when no danger signals are apparent... There are at present no firm rules; each programmer has his own."(Markle, 1967,pps. 122-123)

While in her statement Dr. Markle does not rule out the possibility of the evolution of guidelines, she suggests, as do others, that the complex and personal activity of developmental testing may prove to be unique to each person engaging in it.

We would suggest that some guidelines--- perhaps heuristics rather than algorithms--- are certainly possible. If guidelines for interviewing clients in social work agencies or patients in psychological clinics, and if techniques for carrying out complex needs assessments can be (and have been) developed to the point at which they can be transferred to computer-based systems, we see no reason for developmental testing to remain a totally intuitive activity.

On the second point above, namely, that such guidelines simply are not needed, we suggest that it would seem illogical for instructional developers, nearly all of whom heartily endorse the procedures of testing and revision to so contend. Furthermore, the research literature supports the contention that formative evaluation does make a difference---that instruction is significantly improved when revision is based on developmental testing,as the studies of VanderMeer show (e.g., VanderMeer and Montgomery,1964.)

Current research. As first steps in the evolution of an agreed upon set of developmental testing guidelines tied to empirical support,we plan two further research studies in the coming year(see Note 3.)

We have developed and sent out a questionnaire to some 150 developers (see Appendix C.) The questionnaire is designed to elicit responses on several aspects of testing and revision as well as some overall opinions about both activities. In addition to probing respondents for information about their own develomental testing and revision procedures, we ask each to provide a rationale for using that particular procedure. After pretesting the form with a dozen respondents the final version has been distributed to developers in the United States and Canada. We expect the results to be tabulated by

early summer,1984. The statements or guidelines that make up the items can be directly traceable to statements extracted from the literature; thus each statement within a item is keyed to an author or authors. Space is provided on the form to allow the repondent to add his or her own guidelines or procedures. The data from this study will indicate the degree to which the actual operating procedures of developers correspond to the guidelines suggested in the literature.

In the next phase of our work we will be exploring experimentally the effect of a few major variables in developmental testing (i.e., role of developer, role of student) on the kind and amount of data produced by testing. Future research will experimentally examine some aspects of the revision process.

In conclusion. We have found that the almost universally endorsed activity of developmental testing has not yet been provided with support in the literature in the form of agreed upon guidelines for that activity. We recognize the appropriateness of developing such guidelines concurrently with the development of guidelines for revision. We suggest that research in this area is called for and have undertaken one program of research aimed at providing some empirical support for the critical activities of testing and revision.

NOTES

1.This paper was prepared for presentation at the Annual meeting of the American Educational Research Association, New Orleans, April,1984.

The senior author's address is:

George L. Geis

Centre for Teaching and Learning Services

McGill University

Macdonald-Harrington Building

815 Sherbrooke St. West

Montreal, Quebec, Canada H3A 2K6

2. We opt for the term "developmental testing. This is not because we wish to avoid paying the apparently obligatory obeisance to Michael Scriven as godfather of formative evaluation. We find the term (i.e.,formative evaluation) too broad and ill-defined for the specific activities we are addressing. Similarly we have not chosen Stolovitch's Learner Verification and Revision (see, e.g., Stolovitch,1982) because it implies a continuing process of testing and revision---something which we heartily endorse but think should, for practical purposes, be distinguished from the activities of testing and revising instruction when it is still under development in prototypic form. In favor of "developmental testing" is its clariity of meaning and more precise reference.

3.The research to be carried out June 1984- June 1985 is supported by the Province of Quebec's Programme de formation de chercheurs et d'action concertee (FCAC.)

References

Abedor, A.
Development and validation of a model explicating the formative evaluation process for multi-media self-instructional learning systems. Doctoral dissertation, Michigan State University,Lansing, Michigan, 1971. (ERIC ED 064 327,1972.)

Andrews,D.H. and Goodson,L.A. A comparative analysis of models of instructional design. Journal of Instructional Development. 1980, 3, 4, 2-16.

Baker,E.L. and Alkin, M.C. Formative evaluation of instructional development. AV Communication Review, 21, 4, 1973,pps.389-418.

Davis,R.H., Alexander, L.T., and Yelon,S.L. Learning system design: an approach to the improvement of instruction. New York: McGraw-Hill,1974.

Deterline, W.A. and Lenn, P.D. Coordinated instructional systems. Palo Alto, Calif.: Sound Education,Inc.,1972.

Dick,W. Formative evaluation in instructional development. J. Inst. Dev. 1980,3 (3),3-6.

Dick,W. and Carey,L. The systematic design of instruction. Glenview,Ill: Scott,Foresman and Company,1978.

Gropper,G.L. Diagnosis and revision in the development of instructional materials. Englewood Cliffs,New Jersey: Educational Technology Publications,1975.

Horn, R.E. Developmental Testing. Ann Arbor,Michigan: Center for Programmed Learning for Business,1966.

Johnson, R.B. and Johnson,S.R. Toward individualized learning: a developer's to self instruction. Reading, Mass.:Addison-Wesley Pub. Co., 1975.

Lawson,T.E. Formative instructional product evaluation: Englewood Cliffs, New Jersey: Educational Technology Publications,1974.

Markle, S.M. Empirical testing of programs (chp.V) In Lange, P.C. (Ed.) Programmed instruction, the sixty-sixth yearbook of the National Society for the Study of Education,Part II. Chicago: University of Chicago Press,1967.

Paulson,C.F. Evaluation of instructional systems. In Crawford,J.(Ed.), CORD: National Research Training Manual. Teaching Research Division, Oregon State System of Higher Education,1969.

Performance and Instruction Journal, June,1983, 22, 5.

Robeck, M. D. A study of the revision process in programmed instruction. Unpublished Masters Thesis, University of California: Los Angeles, 1965.

Robinson, T. J. Replicable training in revision techniques. Doctoral dissertation, University of California, L.A., 1972. Dissertation Abstracts International, 33/04-A, p. 1573, 1972. Univ. Microforms E 72-25, 826.

Sanders, J.R. and Cunningham, D.J. A structure for formative evaluation in product development. Rev. of Educ. Res., 43, 1973, 2, 217-236.

Stolovitch,H. Application of the intermediate technology of learner verification and revision (LVR) for adapting international instructional resources to meet local needs. Performance and Instruction. September,1982,21 (7),16-22.

VanderMeer, A. W. and Montgomery, R.E. An investigation of the improvement of educational filmstrips and the derivationof principles relating to the effectiveness of these media. Study No. 3. University Park, Pa., Pennsylvania State University, 1964.

Wager, J.C. One-on-one and small group formative evaluation. Performance and Instruction. 22, 5, June 1983, 5-7.

Weston,C.,Burt,C. and Geis,G. Instructional development: revision procedures. Paper presented at the annual meeting of the American Educational Research Association, New Orleans,1984.

Appendix A

Abedor, A. Development and validation of a model explicating the formative evaluation process for multi-media self-instructional learning systems. (Doctoral Dissertation) Michigan State University, 1971.

Baker, E. L., and Schutz, R. E. Instructional product development. New York: Van Nostrand Reinhold, 1971. Bell, N. and Abedor, A. Developing audio-visual instructional modules for vocational and technical training. Englewood Cliffs, N.J.: Educational Technology Publications, 1977.

Bjerstedt, A. Educational technology: instructional programming and didakometry. New York: McGraw Hill, Inc., 1972.

Davis, R. H., Alexander, L. T., and Yelon, S. L. Learning system design: an approach to the improvement of instruction. New York: McGraw-Hill Book Co., 1974.

Deterline, W. and Lenn, P. Coordinating instructional systems training: a self-study course for teachers. Palo Alto, California: Sound Education Inc., 1972.

Dick, W. and Carey, L. The systematic design of instruction. Glenview, Illinois: Scott, Foresman and Company, 1978.

Friesen, P. Designing instruction. Santa Monica, California: Miller Publishing Company, 1973.

Gropper, G. L. Diagnosis and revision in the development of instructional materials. Englewood Cliffs, New Jersey: Educational Technology Publications, Inc., 1975.

Horn, R. E., Developmental testing. Ann Arbor, Michigan: Center for Programmed Learning for Business, 1966.

Johnson, R. and Johnson, S. Toward individualized learning: a developer's guide to self-instruction. Reading, Massachusetts: Addison-Wesley Publishing Company, 1975.

Lawson, T. Formative instructional product evaluation: instruments and strategies. Englewood Cliffs, New Jersey: Educational Technology Publications, 1974.

Markle, S. Designs for instructional designers. Champaign, Illinois: Stipes Publishing Company, 1978.

Merrill, M. D., Reigeluth, C. M., and Faust, G.W.

The instructional quality profile: a curriculum evaluation and design tool. in O'Neill, Jr., H. F. (Ed.), Procedures for instructional systems development. New York: Academic Press, 1979.

Merrill, D. M. and Tennyson, R. D. Teaching concepts: an instructional design guide. Englewood Cliffs, New Jersey: Educational Technology Publications, Inc., 1977.

Thiagarajan, S., Semmel, D., and Semmel, M. Instructional development for training teachers of exceptional children: a sourcebook. Bloomington, Indianna: Center for Innovation in Teaching the Handicapped, 1974.

Wydra, F. Learner controlled instruction. Englewood Cliffs, New Jersey: Educational Technology Publications, 1980.

Appendix B

1. What is criterion performance?
2. How many test subjects are used during developmental testing?
3. What are the characteristics of the test subject population?
4. What is the role of the test subject during developmental testing and revision?
5. What is the role of the developer during developmental testing and revision?
6. How is data colected during developmental testing?
7. What is revision based on?
8. Testing environment

FIGURE 1
ROLE OF TEST SUBJECT

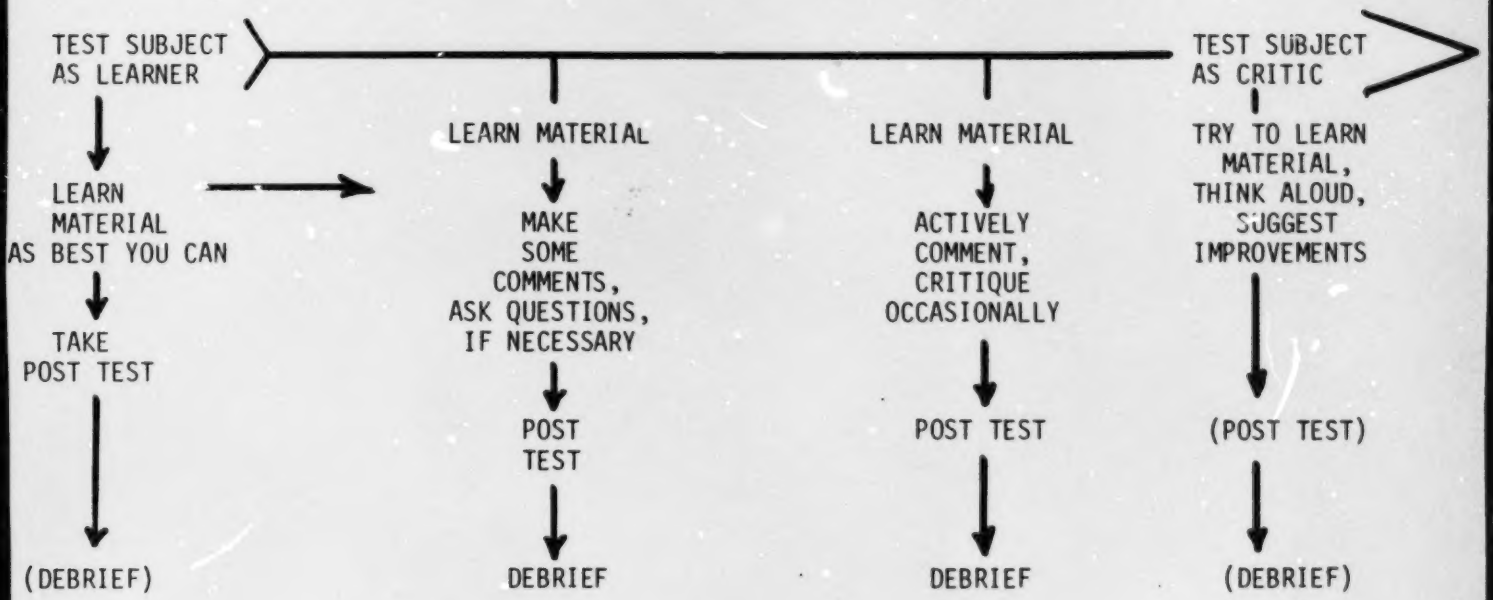
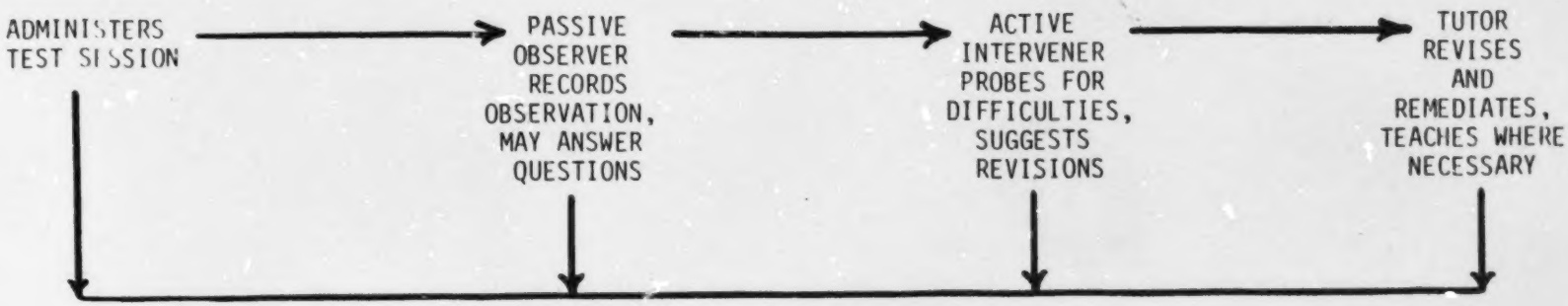


FIGURE 2

ROLE OF DEVELOPER


ADMINISTERS TEST SESSION
PASSIVE OBSERVER RECORDS OBSERVATION, MAY ANSWER QUESTIONS
ACTIVE INTERVENER PROBES FOR DIFFICULTIES, SUGGESTS REVISIONS
TUTOR REVISES AND REMEDIATES, TEACHES WHERE NECESSARY
CONDUCTS DEBRIEFING

Geis, Weston & Burt
Aera, 1984

FIGURE 3

AMOUNT OF PRECISE ADVICE AVAILABLE

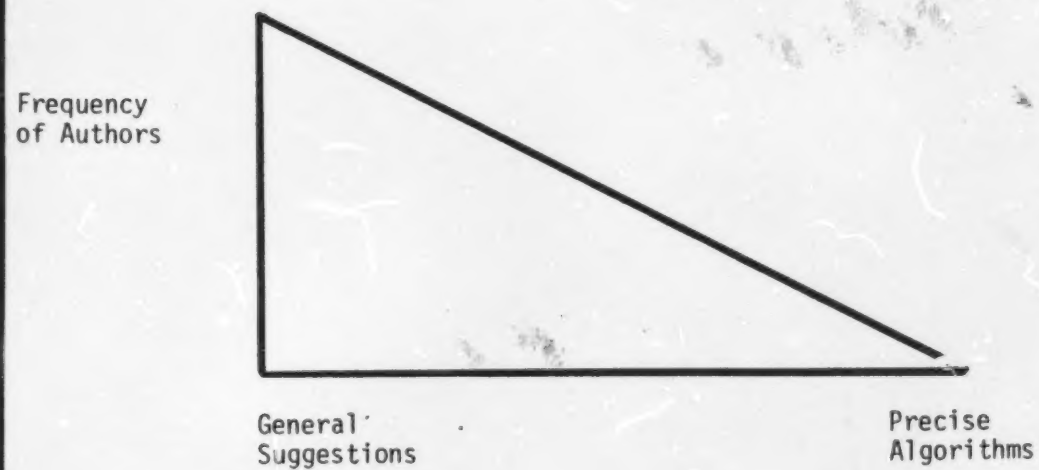


TABLE I

	Criterion Performance	Number of Students	Population Characteristics	Role of Student	Role of Developer	Method Data Collection	Basis of Revision	Testing Environment
ABEDOR	80%/80% on post test; 80% on reactionnaire	9-12 in small group debriefing	varying abilities (SAT scores)	semi-passive during testing; active during debriefing	passive during testing; active during debriefing	post tests, reactionnaire, tape recordings, observation, debriefing comments	post test, reactionnaire, debriefing	cites Horn
BAKER & SCHUTZ	test until performance variability eliminated	not too small; not too large	-	passive learner	summarizes testing data (developer and tester may be different)	depends on type/format of test	criterion test and "Student Response"	-
BELL & ABEDOR	70-80% of students correct on test item; 70% satisfaction on reactionnaire	4-6 in small group debriefing	varying abilities	(see ABEDOR)	(see ABEDOR)	(see ABEDOR)	student revisions (consider cost-benefit); (also: see ABEDOR	-
BJERSTEDT	-	one or more	typical of population; not too timid.	active during testing: thinking aloud individual debriefings; completes pre and post test and reactionnaire	tutors student during testing	pre/post test; reactionnaire; student comments; debriefing.	error analysis; student comments	-
DAVIS, ALEXANDER & YELON	-	-	representation group	-	designs tests, debriefs	achievement tests; reactionnaire	-	-

BEST COPY AVAILABLE

	Criterion Performance	Number of Students	Population Characteristics	Role of Student	Role of Developer	Method Data Collection	Basis of Revision	Testing Environment
Deterline & Lenn	tutor student until meets objective	3 students; one-to-one	representative of target population	semi-passive during testing; active during debriefing	tutors during testing; debriefs; analyzes results	tape-recorded tutoring (if required); notes; probing; post test	revision algorithm; types of errors	-
Dick & Carey	-	2 or more for 1 to 1; small group of 10-20	high, average, low in general ability	active, discusses difficulties	active; revises on the spot	notes, revisions made during testing	errors on test items; student comments, revisions, reactions, developer notes.	-
Friesen	7% error rate on "frame"	one-to-one until success; small group (6); large group (30)	representative of target population	complete materials and tests; responds to developer's probes.	semi-passive during testing; intervenes if required.	notes on testing; analysis of test results.	student testing behaviours and test results.	-
Gropper	——— see BASIS OF REVISION ———						provides numerous decision tables & algorithms implying various rules and strategies.	-
Horn	-	varies from 1-12, usually 3	representative of target population; include high, average, low.	completes material, writes or expresses comments; incl. confidence level.	introduce testing; give tests; times, revises on spot; takes notes; interviews.	(see ROLE OF DEVELOPER)	student comments; test data; on-the-spot revision.	informal; free from distraction.

BEST COPY AVAILABLE

	Criterion Performance	Number of Students	Population Characteristics	Role of Student	Role of Developer	Method Data Collection	Basis of Revision	Testing Environment
Johnson & Johnson	70%/90%	2 or 3	anyone	work through material recording difficulties	revises on spot, if possible.	student revision comment sheet	apparent difficulties	-
Lawson	-	one-on-one; small group	use below average student	completes materials, tests and questionnaires; criticizes	records all diagnostic information; discusses difficulties	numerous forms for internal and external evaluation	expert input, student opinions and comments, test results.	-
Markle	acceptable test data	-	testing reveals critical prerequisites	complete materials and tests	analyzes test data (emphasis on valid criterion test)	test data	test results	-
Merrill/ Reigeluth/ Faust	INTERNAL REVIEW USING INSTRUCTIONAL QUALITY PROFILE							
Merrill/ Tennyson	test results; confidence rating by student	one-on-one, (repeat as often as necessary); then small group (N - 15)	-	go through materials; makes comments; take tests; note confidence one-on-one.	observe student and take notes, clarify points of difficulty	notes by student and developer; on-the-spot revision; pre/ post tests; student confidence ratings.	early revisions based on clinical testing; later revision based on group test data, algorithms for revision.	-

BEST COPY AVAILABLE

	Criterion Performance	Number of Students	Population Characteristics	Role of Student	Role of Developer	Method Data Collection	Basis of Revision	Testing Environment
Thiagarajan/ Semmel/ Semmel	until consistent results are obtained	"qualitative testing" 4 or 5 in 1-to-1 or group; "quantitative testing - small group.	enthusiastic and verbal students	one-on-one: go through materials, think aloud, request assistance; small group: act as learner	one-on-one-monitors, gives assistance when needed; administer and monitor group session.	student verbal/ non-verbal reactions; interviews; pre/post test group data.	on-the-spot changes; test data for group.	-
Wydra	-	-	-	learner; makes comments	passive during testing; "inconspicuous", administers session.	tape record session; student written comments; developer comments (written on material).	notes by student and developer; test data	-

BEST COPY AVAILABLE

Appendix C

A SURVEY-QUESTIONNAIRE ON THE PROCESS OF DEVELOPMENTAL TESTING AND REVISION OF PROTO-TYPE INSTRUCTIONAL MATERIALS

Center for Teaching and Learning Services
McGill University
MacDonald-Harrington Building
815 Sherbrooke St. West
Montreal, Quebec, Canada, H3A 2K6

RESEARCHERS

Dr. G. Geis
Dr. C. Weston
Mr. C. Burt

FIRST-DRAFT VERSION: NOT TO BE CITED WITHOUT PERMISSION FROM THE AUTHORS

BEST COPY AVAILABLE

1. In the developmental testing and revision phase of the design of proto-type instructional materials, which of the following general strategies do you usually employ? PLEASE CHECK, ☑ , THE MOST APPROPRIATE RESPONSE(S).

- ☐ testing with subjects on a one-to-one basis.
- ☐ testing with subjects on a small group basis.
- ☐ revision is based on criterion test data.
- ☐ revision is based on an attitudinal survey.
- ☐ I elicit test subject comments.
- ☐ revision takes place during the testing session.
- ☐ revision takes place after the testing session.
- ☐ OTHER (please outline briefly)_____

_____

_____

_____

_____

2. What aptitudes or levels of performance do you require of test subjects utilized during the developmental testing and revision phase of the instructional design process? PLEASE CHECK, ☑ , THE MOST APPROPRIATE RESPONSE(S).

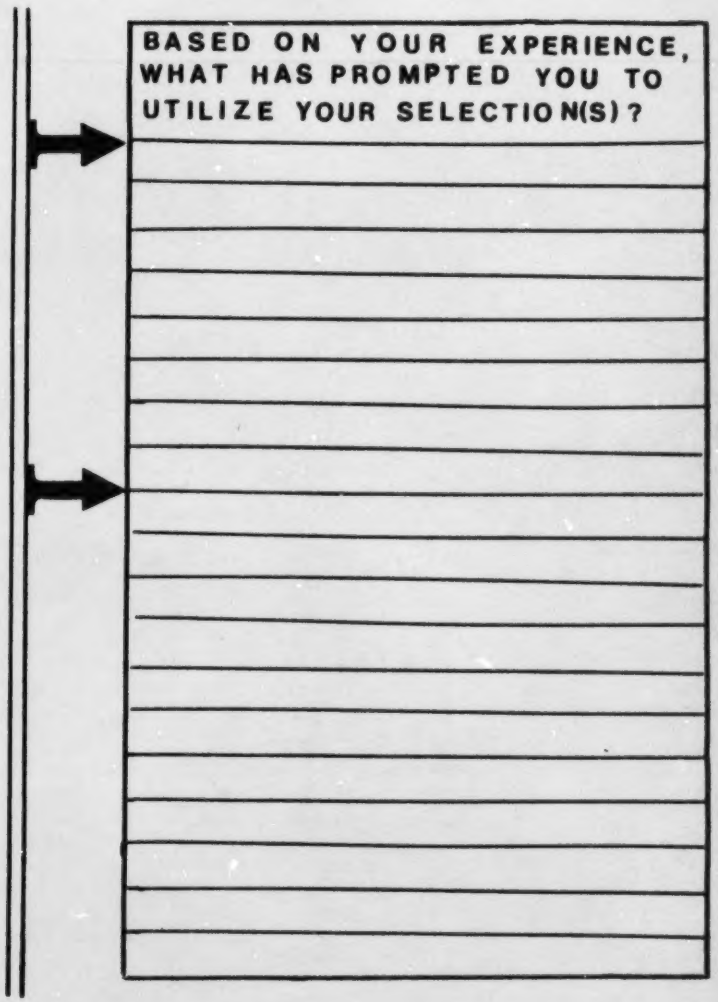
- ☐ high level/aptitude performers only.
- ☐ high and low level/aptitude performers.
- ☐ high, medium, and low level/aptitude performers.
- ☐ enthusiastic and highly verbal test subjects.
- ☐ OTHER (please outline briefly)_____

_____

_____

_____

_____



3. What kind of instrument do you usually use to differentiate student level/aptitude? PLEASE CHECK, ☑ , THE MOST APPROPRIATE RESPONSE(S).

- ☐ I.Q.
- ☐ a standardized aptitude test.
- ☐ past student achievement data (e.g., G.P.A., course grades, etc.)
- ☐ job performance measures.
- ☐ OTHER (please outline briefly)_____

_____

_____

_____

_____

BASED ON YOUR EXPERIENCE, WHAT HAS PROMPTED YOU TO UTILIZE YOUR SELECTION(S)?

4. How many test subjects do you typically involve during a developmental testing and revision session? PLEASE CHECK, ☑ , THE MOST APPROPRIATE RESPONSE(S).

- ☐ 3 or 4 test subjects in one-to-one sessions.
- ☐ 4 to 6 test subjects in one-to-one sessions.
- ☐ 4 to 6 test subjects in small group sessions.
- ☐ 6 to 12 test subjects in small group sessions.
- ☐ OTHER (please outline briefly)_____

_____

_____

_____

_____

_____

BASED ON YOUR EXPERIENCE, WHAT HAS PROMPTED YOU TO UTILIZE YOUR SELECTION(S)?

5. What do you consider the role of the test subject to be during developmental testing? PLEASE CHECK, ☑, THE MOST APPROPRIATE RESPONSE(S).

- ☐ test subject provides pre- and post-test criterion data.
- ☐ test subject provides attitudinal data.
- ☐ test subject comments on problems with the instructional material.
- ☐ test subject comments on problems with the tests.
- ☐ OTHER (please outline briefly)_____

_____

_____

_____

_____

_____

BASED ON YOUR EXPERIENCE, WHAT HAS PROMPTED YOU TO UTILIZE YOUR SELECTION(S)?

6. What do you consider the role of the test subject to be during the revision of the instructional materials? PLEASE CHECK, ☑ , THE MOST APPROPRIATE RESPONSE(S).

- ☐ the role of the test subjects is to learn as evidenced in pre- and post-test data.
- ☐ test subjects comment by verbally indicating problems.
- ☐ test subjects are asked to aggressively determine the causes of their problems.
- ☐ test subjects are asked to determine the causes of their problems only when the instructional developer notices a problem.
- ☐ test subjects comment by writing on the instructional material itself or on a revision sheet.
- ☐ test subjects actually revise instructional material as they work through it.
- ☐ test subjects actually revise instructional material after they work through it.
- ☐ OTHER (please outline briefly)_____

BASED ON YOUR EXPERIENCE, WHAT HAS PROMPTED YOU TO UTILIZE YOUR SELECTION(S)?

7. During a developmental testing session, test subjects are:
PLEASE CHECK, ☑, THE MOST APPROPRIATE RESPONSE(S).

- ☐ not permitted to ask any questions.
- ☐ not permitted to make comments.
- ☐ permitted to but not encouraged to make comments or ask questions.
- ☐ actually encouraged to ask questions or make comments.
- ☐ actually taught by the developer when they run into difficulty.
- ☐ thinking aloud and requesting assistance when required.
- ☐ recording problems on a revision sheet and/or on the actual instructional materials as they work through the material.
- ☐ OTHER (please outline briefly)_____

_____

_____

_____

_____

_____

BASED ON YOUR EXPERIENCE, WHAT HAS PROMPTED YOU TO UTILIZE YOUR SELECTION(S)?

8. What do you consider the role of the developer to be during developmental testing? PLEASE CHECK, ☑, THE MOST APPROPRIATE RESPONSE(S).

- ☐ the developer waits, watches, makes notes, and keeps quiet while the test subject works.
- ☐ if a problem arises, the developer asks "why", pinpoints the problem, and gives just enough help to get the test subject back on track.
- ☐ the developer personally answers test subject's questions in a tutorial fashion.
- ☐ the developer notices gaps, predicts confusions, senses response failures, and re-writes items on the spot.
- ☐ OTHER (please outline briefly)_____

_____

_____

_____

_____

_____

BASED ON YOUR EXPERIENCE, WHAT HAS PROMPTED YOU TO UTILIZE YOUR SELECTION(S)?

9. What do you consider the role of the developer to be during the revision of proto-type instructional materials? PLEASE CHECK, ☑, THE MOST APPROPRIATE RESPONSE(S).

- ☐ the developer revises on the spot after the test subject has commented.
- ☐ the developer debriefs a small group of test subjects based on criterion test data, attitudinal survey, and elicited test subject comments. Revision is then based on these debriefing results.
- ☐ the developer revises based on data generated from teaching the instructional material to the test subject in a tutorial fashion.
- ☐ the developer revises based on a comment sheet that the test subject has used to point out problem areas.
- ☐ the developer revises on the basis of criterion test data and attitudinal survey after each one-to-one session.
- ☐ OTHER (please outline briefly)_____

BASED ON YOUR EXPERIENCE, WHAT HAS PROMPTED YOU TO UTILIZE YOUR SELECTION(S)?

10. When do you stop developmental testing? PLEASE CHECK, ☑ , THE MOST APPROPRIATE RESPONSE(S).

- ☐ after a pre-determined number of one-to-one sessions.
- ☐ after having debriefed students from one-to-one sessions.
- ☐ after having debriefed students in a small group setting.
- ☐ after criterion test results begin indicating a "success" factor.
- ☐ when a previously set criterion level of achievement and attitudinal response is being met during one-to-one testing.
- ☐ when a previously set criterion level of achievement and attitudinal response is being met from small group debriefing sessions.
- ☐ OTHER (please outline briefly)_____

BASED ON YOUR EXPERIENCE, WHAT HAS PROMPTED YOU TO UTILIZE YOUR SELECTION(S)?

11. What is the most common criterion performance standard that you set to determine the cessation of revision during the proto-type testing phase? PLEASE CHECK, ☑ , THE MOST APPROPRIATE RESPONSE(S).

- ☐ I usually set no specified criteria for test subject performance.
- ☐ when less than 7% of the test subjects are having problems with the material.
- ☐ 4 or 5 sessions (individual or small group) until consistent results are obtained.
- ☐ not more than 20% unsatisfactory responses on the attitudinal survey.
- ☐ 70% of the test subject group achieving 70% or higher on the post-test.
- ☐ 80% of the test subject group achieving 80% or higher on the post-test.
- ☐ OTHER (please outline briefly)_____

_____

_____

_____

_____

_____

BASED ON YOUR EXPERIENCE, WHAT HAS PROMPTED YOU TO UTILIZE YOUR SELECTION(S)?

12. How do you obtain your data during developmental testing? PLEASE CHECK, ☑, THE MOST APPROPRIATE RESPONSE(S). AFTER HAVING CHECKED OFF YOUR RESPONSE, PLEASE INDICATE BY PLACING A STAR, ☐✡, NEXT TO THE ITEM(S) PROVIDING THE MOST USEFUL DATA FOR REVISION.

- ☐ from criterion pre- and post-tests.
- ☐ from an attitudinal survey.
- ☐ from the probing (through interviewing techniques) of test subjects.
- ☐ by recording (by note-taking, audio tape, or video tape) information.
- ☐ from inferences based on observed test subject behavior.
- ☐ by debriefing test subjects.
- ☐ from test subject generated comments and notations written on the instructional materials.
- ☐ from test subject comments placed on a revision comment sheet.
- ☐ from the notes that were taken on examples, illustrations, information added, and changes to sequences made while the test subject was tutored through the instructional material.
- ☐ from the comments that the test subjects have made verbally.
- ☐ OTHER (please outline briefly)_____

BASED ON YOUR EXPERIENCE, WHAT HAS PROMPTED YOU TO UTILIZE YOUR SELECTION(S)?

13. How do you analyze the data gathered from developmental testing sessions? PLEASE CHECK, ☑, THE MOST APPROPRIATE RESPONSE(S).

- ☐ the extent of the analysis depends upon the resources available (e.g., budget and time constraints).
- ☐ by comparing the cumulative effects of revisions based on a number of test sessions (e.g., by testing out the revisions generated from a session on the next test subject(s).
- ☐ by debriefing test subjects after their completion of the material, thus providing tallied test results, attitudinal data, and test subject generated comments. The test subjects and the developer cooperatively analyze the data.
- ☐ by comparing pre- and post-test data with a pre-set standard.
- ☐ with strong emphasis placed on error analysis to find out why individual test subjects make errors.
- ☐ OTHER (please outline briefly)_____

_____

_____

_____

_____

_____

BASED ON YOUR EXPERIENCE, WHAT HAS PROMPTED YOU TO UTILIZE YOUR SELECTION(S)?

14. For purposes of revision after you have obtained information from developmental testing, do you: PLEASE CHECK, ☑, THE MOST APPROPRIATE RESPONSE(S).

- ☐ follow a prescribed algorithm or heuristic, e.g., certain kinds of errors dictate certain changes.
- ☐ revise on the basis of "obvious", self-evident weaknesses within the instructional material.
- ☐ hard to say; it is intuitive.
- ☐ OTHER (please outline briefly)_____

_____

_____

_____

_____

_____

15. We recognize that many instructional developers have evolved sophisticated procedures over the years, for developmental testing and revision. We are interested in learning how explicit and teachable your procedures are, e.g., can they be taught to others? PLEASE CHECK, ☑, THE MOST APPROPRIATE RESPONSE(S).

A. With regard to developmental testing?

☐ my procedures are well documented in the literature or in handouts, and could be/are taught to others.

☐ my procedures could easily be documented and taught to others but have not been.

☐ it would be difficult for my procedures to be documented or taught to others.

☐ my procedures can not be (or are not likely to be) documented or generalized because of varying local conditions, e.g., each project, test subject, context, etc., is different and requires a different strategy.

☐ OTHER (please outline briefly)_____

_____

_____

_____

_____

_____

B. With regard to revision?

- ☐ my procedures are well documented in the literature or in handouts, and could be/are taught to others.
- ☐ my procedures could easily be documented and taught to others but have not been.
- ☐ it would be difficult for my procedures to be documented or taught to others.
- ☐ my procedures can not be (or are not likely to be) documented or generalized because of varying local conditions, e.g., each project, test subject, context, etc., is different and requires a different strategy.
- ☐ OTHER (please outline briefly)_____

_____

_____

_____

_____

_____

➘ We would very much appreciate your enclosing or providing a reference to any documentation on developmental testing or revision that you have generated and/or use.

- ☐ ENCLOSING
- ☐ NOT ENCLOSING

16. What percent of your time, roughly, is devoted to each of the following activities?

1. _____% Management of an instructional design unit.
2. _____% Design and production of instructional products and materials.
3. _____% Teaching instructional design within an academic setting.
4. _____% Teaching instructional design within a non-academic setting.
5. _____% Consulting on instructional design projects.
6. _____% Researching within the field of instructional design.
7. _____% Activities related to the profession, e.g., editor of journals, officer in a professional association, etc.
8. _____% Presenting and publishing within the field of instructional design.

17. Have you published or produced books, articles, video tapes, or slide shows related to instructional design, within the past three years.

☐ YES

☐ NO

18. In what kind of an institution do you work?

- ☐ 1. Academic
- ☐ 2. Business or Industry
- ☐ 3. Military
- ☐ 4. OTHER (please outline briefly)_____

_____

_____

_____

_____

_____

19. How much did each of the following contribute to you becoming an Instructional Developer? PLEASE GIVE A ROUGH PERCENTAGE.

1. _____% Academic Training
2. _____% On the Job Training
3. _____% Experience and Self-Teaching
4. _____% Workshops, Seminars, and Conferences
5. _____% OTHER (please outline briefly)_____

_____

_____

20. This page is at your disposal to indicate changes or suggestions that you believe will help us improve the survey and attack the issue at hand. It can also be used to summarize your overall reaction to the survey. Thank you for your time and consideration.